

# System and method for sending and receiving a video as a slide show over a computer network

**Publication number: AU6672298**

**Publication date:** 1998-09-09

**Inventor:** COLBY KENNETH W; KENNER BRIAN; WEATHERSBY  
GUY P; BROWNELL LONNIE J; FLYNN PETER K

**Applicant:** INTERVU INC

**Classification:**

- International: **H04N7/26; H04N7/52; H04N7/26; H04N7/52; (IPC1-7):**  
**H04N7/26**

- European: H04N7/26E6; H04N7/26T; H04N7/52A

Application number: AU19980066722D 19980225

Priority number(s): US19970806093 19970225; WO1998US03904  
19980225

**Also published as:**

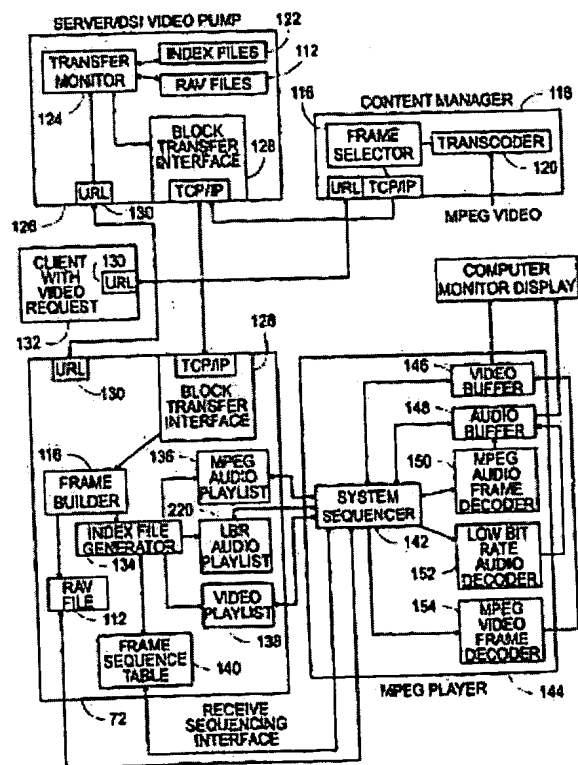
 WO9837699 (A1)

**Report a data error here**

Abstract not available for AU6672298

Abstract of corresponding document: **WO9837699**

A system and method for encoding and decoding digitized audio/video files prepares a slide show of still images and a low bit rate audio stream which can be downloaded in real time over a typical connection to a computer network. The quality of audio/video file is subsequently improved by downloading in successive passes the remaining video frames, which are restored to their original order, and the original high-quality audio content.



Data supplied from the **esp@cenet** database - Worldwide

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

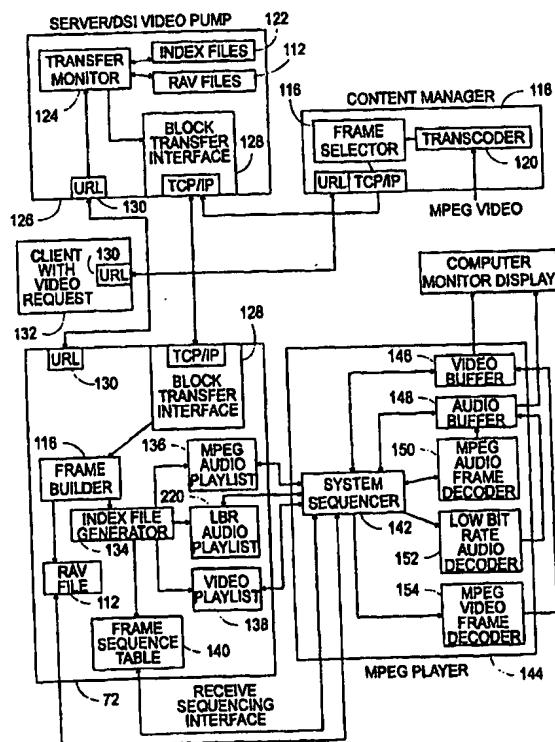
(51) International Patent Classification <sup>6</sup> : <b>H04N 7/26</b>		<b>A1</b>	(11) International Publication Number: <b>WO 98/37699</b>
			(43) International Publication Date: 27 August 1998 (27.08.98)
(21) International Application Number: PCT/US98/03904		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 25 February 1998 (25.02.98)			
(30) Priority Data: 08/806,093 25 February 1997 (25.02.97) US			
(71) Applicant: INTERVU, INC. [US/US]; 201 Lomas Santa Fe Drive, Solana Beach, CA 92075 (US).			
(72) Inventors: COLBY, Kenneth, W.; 12707 Gibraltar Drive, San Diego, CA 92128 (US). KENNER, Brian; 1403 Walnut Creek Drive, Encinitas, CA 92024 (US). WEATHERSBY, Guy, P.; 8674 Perseus Road, San Diego, CA 92125 (US). BROWNELL, Lonnie, J.; 826 Birchview Drive, Encinitas, CA 92024 (US). FLYNN, Peter, K.; 14742 Via Abertura, P.O. Box 8204, Rancho Santa Fe, CA 92067 (US).			
(74) Agents: WIXON, Clarke, A. et al.; Darby & Darby P.C., 32nd floor, 707 Wilshire Boulevard, Los Angeles, CA 90017 (US).			

**Published***With international search report.**Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*

(54) Title: SYSTEM AND METHOD FOR SENDING AND RECEIVING A VIDEO AS A SLIDE SHOW OVER A COMPUTER NETWORK

**(57) Abstract**

A system and method for encoding and decoding digitized audio/video files prepares a slide show of still images and a low bit rate audio stream which can be downloaded in real time over a typical connection to a computer network. The quality of audio/video file is subsequently improved by downloading in successive passes the remaining video frames, which are restored to their original order, and the original high-quality audio content.



**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

**SYSTEM AND METHOD FOR SENDING AND RECEIVING A VIDEO  
AS A SLIDE SHOW OVER A COMPUTER NETWORK**

The invention relates to a system and method whereby a digitized audio-  
5 video file is reconfigured and downloaded over a computer network to a user  
terminal in successive passes of data, so that during or after each pass, the user can  
see and hear the audio-video file with increasing quality. In a preferred  
embodiment, the audio-video file can be viewed as a high quality slide show with  
low bit rate audio during the download process and replayed as a video with full  
10 audio after completing the download process.

**BACKGROUND OF THE INVENTION**

Video data has extremely high storage and bandwidth requirements. In order  
15 to reduce the bandwidth required to transmit video data, digitized video files can be  
compressed to reduce the data comprising the video file. During the process of  
video compression, video information is deleted that would be imperceptible to the  
human eye. As more video data is deleted the size of the video file decreases and  
the bandwidth required to deliver the video file is reduced. A variety of methods  
20 and protocols exist for compressing digitized video files and are well known in the  
art.

MPEG (Motion Pictures Experts Group) is regarded by many as the standard  
for digital video compression. Videos produced in the MPEG format and played at  
25 a rate of 24 frames per second provide high quality, high resolution video and high  
quality audio.

MPEG video files, like other compressed video files, are still rather large  
compared to smaller text and graphic files, and can take from several minutes to  
30 hours of constant data flow to download. High capacity host/client architecture  
capable of high storage and transmission rates is required to transmit and receive  
this data error-free without corruption or loss of data. In a distributed computer

network such as the Internet, it is difficult, if not impossible, to provide a host/client architecture which has the capacity for accurate, sustained, high speed transmission of large audio/video files.

5 Even where the capacity of the network distribution system is improved to permit more data transfer, a bottleneck typically occurs at the user modem which establishes the connection between the user and the network. A typical user modem only receives data at a rate of 28.8 kilobits per second. A 30 second MPEG video can take 5 minutes or more to download over a 28.8 modem. Because the data is  
10 often transferred from afar, many factors can cause the loss of parts or all of a transmission, thus slowing the receipt as re-transmission of the lost data occurs.

Real time video delivery has even more specific and stringent transfer and display timing requirements. In this case, the user wants to be able to view the  
15 video at the user terminal while the video data is being downloaded. In order to do this, the line between the user terminal and the server must have enough bandwidth to accommodate a steady stream of data comprising all the information necessary for playing the video. If the bandwidth is not available, the data stream will be delayed during the download and there will be insufficient data available at the user  
20 terminal to play back the video in real time, as it was originally encoded. As a result, the user will observe interruptions and delays in the video and audio content.

One attempt to improve real time video delivery has been to further compress the video. To accomplish this, some video content providers compress  
25 the video data by encoding at a slower frame rate of 6-7 frames per second (fps) and encoding the audio data at a lower bit rate, thereby deleting large portions of content. The resulting video has poor quality and very choppy motion and the sound quality is poor. The video and audio data which is deleted during this compression process is permanently lost. Therefore, even if the download is  
30 successful, the quality of the video cannot be improved; it will look and sound just as poor on subsequent replays. Even at this reduced size, the video may consist of more data than can be transmitted at the necessary viewing speed (in real time) over

a 28.8 kbaud modem, so that picture and sound quality is further degraded when the user views it.

Another solution involves a compression format wherein data can be added  
5 to a video file during transmission to progressively improve the image. As the  
video file is being downloaded, the content server is continuously testing the  
bandwidth of the network link to the user and making decisions on a frame-by-frame  
basis whether to pass more or less data to the user. As more bandwidth becomes  
available, more data can be passed down and the quality of the video image and  
10 audio is improved. Like the previous example, the video data is lost and cannot be  
recovered once the video file is downloaded. The resulting video is of uneven  
quality, and subsequent replays will look and sound the same.

Neither solution provides a means to transmit meaningful and entertaining  
15 audio/video data to a user in real time that gives the user the option to replay the  
video in its original format, i.e., a high quality video with high quality sound. The  
invention solves this problem by providing a method and system whereby a digitized  
audio-video file can be reconfigured and downloaded over a computer network to a  
user terminal where it can be viewed as a high quality video slide show with low bit  
20 rate audio during the download process and replayed as a full-motion video with  
high quality audio after completing the download process.

### **SUMMARY OF THE INVENTION**

25 In a first embodiment, the audio portion of an original audio-video (AV) file  
is compressed into a low bit rate (LBR) audio data stream by means known in the  
art. The order of the individual frames comprising the original video data stream is  
then rearranged. In a first pass, a frame selector module is used to select individual  
video frames from among all the frames comprising the original video data stream.  
30 These frames will be stored at the front end of a reconfigured AV file along with the  
LBR audio stream. In subsequent passes, the remaining video frames are selected.

The video frame data, LBR audio data stream and audio data stream of the original AV file are then assembled as an AV file having a selectively reordered download sequence and stored for delivery at a server site. When a video clip is requested by a client, the server downloads the video data to the client according to the selectively reordered sequence. As the "front-loaded" portion of the new AV file, is downloaded, the client is able to view a comprehensive audio/video slide show representative of the whole video.

The "front-loaded" portion of the new AV file comprising the slide show is many magnitudes in size smaller than the original AV file (Fig. 1). Thus, even when the bandwidth available for transmission is limited, as is the case with a 28.8 kbaud modem, a high quality video slide show with audio can still be displayed during the download process because the data stream required to support the slide show and compressed audio is much smaller.

Once the slide show frames and LBR audio portion of the new AV file have been downloaded, the remaining video frames and the original audio data stream are downloaded in stages. The client software displays the front loaded data as a slide show during the download process and then resequences the front-loaded data and remaining video frames into the original order. This makes it possible for the client's player to replay portions of the video clip as a low frame rate video during download. If all of the AV data is downloaded, the client software can display the video in its original format and speed with the originally recorded audio quality.

In a second embodiment, the audio portion of an original AV file is highly compressed into an LBR audio data stream by means known in the art. A reconfigured AV file is created consisting of the LBR audio data stream, the original audio data stream and a resequenced video data stream. The frame selector module is used to determine different download orders of video frame data for a variety of given connection speeds. A corresponding index file is created for each download order. The index file records both the download order and information for locating the video data in the new AV file for reassembly in the original order. A frame

sequencing interface (FSI) is responsible for delivering AV files from the server to the client. The FSI, among other functions, reads the index file that matches the client's connection speed and downloads the video frame data to the client according to the order recorded on the index file.

5

In both embodiments, the client software downloads the file until the entire AV file is delivered or the user discontinues the download. As each pass of video data is downloaded, the client software reshuffles the data into its original temporal order making it possible for the client to display the video data with progressively  
10 improved quality. Regardless of the number of frames downloaded, each frame is displayed with the full quality of the originally recorded video file. If all the video data is downloaded the video can be displayed in its originally recorded condition with high quality audio.

15

In both embodiments, the user has the option to stop transmission of a reconfigured AV file at any point. The user can elect, for example, to see only the first frame of the video, to view part or all of a slide show with LBR audio, to view a high quality video with LBR audio, or to view a progressively higher quality video with LBR or originally recorded sound. Thus, the user does not have to use up  
20 valuable bandwidth or time waiting for or viewing video content that does not significantly enhance the viewing experience.

In one embodiment, the client software is configured to permit the full download to occur in the background so the user can perform other operations  
25 during the download process. Once the video is completely downloaded, the user can be signaled, and can replay the high quality video. The client software can also interrupt, delay, and later resume the download process when it senses competition for the communication interface.

30

### **BRIEF DESCRIPTION OF THE DRAWINGS**

Fig. 1 is graph comparing the size (in bytes) of an MPEG video file, a low



frame rate video with low bit rate audio, and a video slide show with low bit rate audio;

Fig. 2 is a block diagram representative of a standard MPEG audio/video decoder;

5 Fig. 3 is a block diagram of a video delivery system according to the invention;

Fig. 4 is a flowchart illustrating the operation of a transcoder module according to Fig. 3;

Fig. 5 is a flowchart illustrating the operation of a frame selector module  
10 according to Fig. 3; and

Fig. 6 is a flowchart illustrating the operation of the video delivery system of Fig. 3.

### **DEFINITIONS**

15

The terms used herein have their ordinary meaning in the art, and in addition, specific terms set forth below have the meanings given.

Slide Show. A sequence of visual images or frames presented as a condensed or  
20 slow-motion version of a video presentation or clip. In an embodiment of the invention, a slide show comprises a sequence of video frames taken from an original full motion audio/video data file, rearranged and adjusted in timing and sequence so as to make an attractive and synchronized presentation. A slide show may be presented with or without accompanying audio content.

25

Video Clip. A video clip is a sequence, of any length, of images, with or without audio content (sound), defining a moving picture or animation.

Audio/Video Data File. An audio/video data file is a digitized computer file  
30 representative of a video clip. The audio/video data file can be in any machine readable format and can be compressed, or reduced in size, by any of several known compression techniques, such as MPEG.

Video Data Stream. A video data stream is that portion of an audio/video data file attributable to the storage of visual images. A video data stream typically comprises at least one sequence of video frames, in presentation or viewing order, or indexed to represent a viewing order. Other possible portions of an audio/video data file

5 include an audio data stream and a system stream, such as a timing stream or an index representative of a viewing order.

Audio Data Stream. An audio data stream is that portion of an audio/video data file attributable to the storage of audio content. An audio data stream may be made up

10 of a sequence of audio frames.

Video Frame. A video frame is a single static image taken from a video clip. A sequence of video frames, viewed in fast succession, provides an illusion of motion.

15 Audio Frame. An audio frame is a time-divided portion of an audio data stream. Audio frames typically are used for simplicity in handling and processing audio data streams; there is no necessary relationship between individual audio frames and individual video frames. Moreover, individual audio frames may vary in length.

20 Reconfigured Audio/Video (RAV) File. An RAV file is produced from an audio/video data file, which may be referred to as an original or source file, and includes a video data stream having video frames in a different presentation or viewing order than the original audio/video data file. An RAV file may have one or more video data streams and one or more audio data streams, one of which may be

25 LBR audio. An RAV file may be produced or displayed in one or more passes, and may have less than, more than, or the same audio and video information as the original audio/video data file.

Presentation Order. A presentation order is an order, or sequence, in which audio

30 or video frames are stored in an audio/video data file. In certain compression schemes, such as MPEG, the presentation order of certain video frames may differ from the viewing order, as certain video frames are decoded based on information in

other video frames which have not yet been displayed.

Viewing Order. A viewing order is an order, or sequence, in which audio or video frames are displayed. Viewing order may differ from presentation order.

5

Low Bit Rate (LBR) Audio. LBR audio is highly-compressed sound information derived from the audio content of an original audio/video data file. In an embodiment of the invention, LBR audio frames are interleaved with video frames comprising a slide show, so that both the slide show video frames and the LBR audio frames can be downloaded simultaneously and displayed in real-time; the original (non-LBR) audio data stream can be downloaded at a later time.

Low Frame Rate Video. A low frame rate video is a slow motion or reduced-quality version of an original video clip. An audio/video data file representing a low frame rate video includes a subset of the video frames included in the original audio/video data file.

Transcoder Module. In an embodiment of the invention, a transcoder module is a combination of computer hardware and software that decodes an audio/video data file, extracts its video stream and audio stream, and optionally compresses the audio stream into LBR audio.

Frame Selector Module. In an embodiment of the invention, a frame selector module is a combination of computer hardware and software that allows certain video frames to be selected from an audio/video data file for use in a slide show or low frame rate video. Information taken from the selection process is used to generate an RAV file or an index file.

Frame Sequencing Interface (FSI). In an embodiment of the invention, an FSI is used to generate a second version of an RAV file, having a different viewing order or presentation order, from a first RAV file and an index file. The second version can then be transmitted over a communication link having different properties than

the one for which the first RAV file was created.

User Terminal. A computer system capable of displaying audio/video data files. A user terminal may be coupled to a communications network.

5

Server. A computer system coupled to a communications network, capable of transmitting (downloading) stored information to another computer system coupled to the network.

10

### **DETAILED DESCRIPTION OF THE INVENTION**

For purposes of definition, the term video data, as used herein, can mean both video frame data and audio frame data or just video frame data. To display video data means to process an audio-video file in a computer so video images are displayed on the computer monitor and corresponding audio is broadcast on the computer speakers. The term playback or played back has the same meaning as display.

The invention as described below and in each of the following examples is discussed in terms of its application to the delivery of video data in the MPEG format, but the scope of the invention is not limited to the MPEG format or to the examples given. MPEG is one protocol for compression of digitized video. There are a number of compression protocols which are used to reduce the size of an AV file, i.e., JPEG, H261, Indeo, Cinepak, AVI, Quicktime, TrueMotion and Wavelet. The invention can easily be adapted by one skilled in the art to reconfigure video data compressed by any of these methods, and such adaptations are within the scope of the invention.

Regardless of the compression protocol used, the corresponding AV file would comprise an original audio data stream, an original video data stream and a user stream containing information related to the synchronization and playback of the audio/video streams. The video data stream consists of encoded information for

q

video frames comprising all of the picture information for a given video. The video frames are arranged in a preselected order so that when they are processed by a video player at a certain speed (frames per second) a full-motion video can be displayed.

5

In the MPEG format, a discrete cosine transform compression algorithm is used to identify and delete redundant video information both between frames and within an individual frame. The video stream of an MPEG movie comprises a series of video frames flanked by a header sequence and an end-of-sequence code.

10 Much of the information in a frame within a video sequence is similar to information in the previous or subsequent frame. The MPEG standard takes advantage of this temporal redundancy by representing some frames in terms of their differences from other (reference) frames.

15 The MPEG standard specifically defines three types of frames: intra, predicted, and bidirectional. Intra (I) frames, are coded using only information present in the frame itself and are present at unpredictable points within the sequential frames of compressed video data. Predicted (P) frames are coded with respect to the nearest previous I or P frame. Bidirectional (B) frames are frames  
20 that use both a past and future frame as a reference. I and P frames both serve as reference frames for B frames. B frames are never used as a reference.

The frequency and location of I frames is based on the need for random accessibility and the location of scene cuts in the video sequence. Where random  
25 access is important, I frames are typically used two times a second. The MPEG encoder reorders the sequence of frames in the video stream to present frames to the decoder in the most efficient sequence. In particular, the I or P reference frames needed to reconstruct B frames are sent before the associated B frames.

30 The MPEG audio stream is similar to the MPEG video stream in that it contains an audio header sequence and one or more audio frames. It should be noted that individual audio frames do not necessarily correspond to individual video  
10

frames. Audio frames are simply "packetized" versions of the audio data, that is, the audio data stream divided into frames by any convenient or useful means. For example, a particular audio compression scheme used to create LBR audio might create frames of substantially equal size, but unequal duration. In contrast, video  
5 frames typically have substantially equal duration but unequal size (in particular, I frames are typically larger than P and B frames).

The timing mechanism that ensures synchronization of audio and video includes two parameters: a system clock (SC) and presentation time stamps (PTS).  
10 The values for these timing mechanisms are coded in the MPEG bitstream. PTS are samples of the system clock that are associated with an individual video frame or audio frame. The PTS indicates the order and timing in which the video frame is to be displayed or the starting playback time for the audio frame.

15 The MPEG AV file consists of both a compression layer and a system layer. The audio and video data streams comprise the compression layer. The system layer contains timing and other information needed to demultiplex the audio and video data streams and to synchronize audio and video during playback.

20 Fig. 2 shows a generalized decoding system for MPEG videos. The system decoder is responsible for extracting the timing information from the MPEG system stream and sending it to the other system components. The system decoder also demultiplexes the video and audio streams from the system stream and sends the data to the appropriate audio or video decoder. Chapter 10 of Video Demystified by  
25 Keith Jack, High Tech Publications, 1996, provides a file format for implementing an MPEG video player that is incorporated by reference and can be adapted for use in the video delivery system described herein.

### **EXAMPLE ONE**

30

A preferred embodiment of the video delivery system allows a user to download a video clip in four passes, the first of which occurs in real time. The  
11

system and method according to which the video delivery is performed is discussed in detail below.

With reference to Fig. 3, a reconfigured AV (RAV) file 112 is created from an MPEG video and stored at a server site 126 on the Internet. A client 132 at a user terminal builds a video request in the form of a URL 130 containing the address of the stored file. The client transmits the URL to the server 126. A connection is made between the client and the server and the server downloads the file to the user terminal (receive sequencing interface) 72 in its precoded order. The user terminal initially processes and displays the slide show data in the order it is received, as it is being received. As additional data is downloaded, it is reshuffled with the slide show data in original temporal order making it possible to replay the video with progressively enhanced quality.

#### 15 Transcoder Module

In Fig. 3, a transcoder module 120 is shown as a component of the content manager 118 of the video delivery system. As will be discussed below, the transcoder module 120 is used in the video delivery system to create an LBR audio data stream and prepare an MPEG video file for resequencing. Accordingly, the transcoder module 120 is used in place of the system decoder of a standard MPEG player (Fig. 2) and performs a similar function.

With reference to Fig. 4, the operations performed by the transcoder module are shown. The transcoder module 120 is used to separate the compression layer of the MPEG file from an original system layer 20. The original system layer is discarded 22 and the transcoder module 120 then disassembles the remaining compression layer into pure MPEG video and MPEG audio data streams 32 and 24, respectively. The data streams 32 and 24 consist of sequential streams of bytes or characters.

The transcoder module 120 compresses the MPEG audio data stream 26

using standard audio compression techniques such as GSM (Global System For Mobile Telecommunications, an international standard for audio compression) to produce a LBR audio data stream which requires transmission bandwidth of approximately 13,000 bits per second or less. The transcoder module 120 also associates 28 a copy of the corresponding PTS with each LBR audio frame indicating the display order of the audio data. Both the original MPEG audio component and the LBR audio component are retained for incorporation into the RAV file.

10 The transcoder module 120, using markers embedded in the MPEG video data streams, locates all of the pure MPEG data necessary to construct a single video frame 34 and encodes that data 36 in an information block (see Table A). Each audio frame in the original MPEG and LBR audio component is also encoded as information blocks 30 and 40. Each block comprises one byte of block ID  
15 representative of the block type, followed by four bytes of block length, followed by the individual block data.

**Table A**  
**(Information Blocks)**

20	Block Type
	Block Length
25	Block Data
	Next Block Type
	Block Length
30	Block Data

35

The file block types are: slide show file header block, I frame block, P frame block, B frame block, video sequence header block, end of video file block, GSM (LBR) audio frame block, and MPEG (high quality) audio frame block. The layout of each type of block is shown in Table B.



**TABLE B****5 Slide Show Header Block**

Variable	Description	Size
Item Type	Block Type (1)	Unsigned Char
Size	Size of Header + data	Long
10 Video Blocks	Number of Video Items	Unsigned Short
LBR Audio Blocks	Number of LBR Audio Blocks	Unsigned Short
LBR Frames Per Block	Number of 33 ms samples in a Block	Short
MPEG Audio Blocks	Number of MPEG items	Unsigned Short
15 Audio Start Time	Offset from start of Video	Long

**I, P and B Frame Blocks**

Variable	Description	Size
Item Type	Block Type (101,102,103)	Unsigned Char
Size	Size of Header + Data	Unsigned Long
Frame Number	Frame play sequence	Unsigned Short
25 Decode Frame Number	Frame decode sequence	Unsigned Short
Transition	Slide to Slide Transition	Short
MPEG Location	Location in MPEG File	Long
PTS	Presentation Time Stamp	Long
Slide PTS	Slide Show PTS	Long
30 VIDEO DATA FOR THE GIVEN FRAME		

**Sequence Header, End of Video File, LBR Audio, and MPEG Audio Blocks**

Variable	Description	Size
40 Item Type	Block Type (103,104,201,202)	Unsigned Char
Size	Size of Header + Data	Unsigned Long
45 DATA FOR THE GIVEN TYPE		

As the audio and video frame data is converted to information blocks, the

blocks are stored in temporary files which retain the data in its original stream order 38, 42, and 44. As these files are created, a temporary index file is generated which records information indicating in which files the sequential audio and video information blocks are located 46. Finally, the index tables and data stream  
5 information are forwarded to a frame selector module 48, as will be discussed in detail below.

### Frame Selector Module

10 The content manager 118 in Fig. 3 also includes a frame selector module 116. The frame selector module is used to select the video data in successive passes for slide show and download sequencing, and thus to encode the RAV and index files. The operations performed by the frame selector module 116 are shown in Fig. 5. The frame selector module 116 is used to select and assemble the data that  
15 will be used to build the RAV file 112. The frame selector module 116 picks the video frame data in successive passes using the index information to choose and locate the respective information blocks. In a first pass, the frame selector 116 picks certain I frame blocks. The chosen I frames are intended to provide a comprehensive "slide show" sampling of the entire video. In a preferred  
20 embodiment, I frames are chosen at a rate no greater than approximately one frame every two seconds. Where an exemplary MPEG file contains two I frames per second, every fourth I frame would be chosen.

The frames that appear in the first pass are chosen as follows; the average bit  
25 size of an I frame is computed 50. The target delivery bandwidth (for example, 28,800 bits per second) is multiplied by a typical usage factor (such as 70%) to give a predicted available bandwidth. The amount of bandwidth needed for the LBR audio is subtracted from the predicted available bandwidth to give the available  
video bandwidth in bits per second (this assures that there is always sufficient  
30 bandwidth to transmit the LBR audio error-free in real time).

The average bit size of an I frame is divided by the available video

bandwidth to give the time needed to download the slide. In a preferred embodiment, this value is used as the interval between slides, unless the number is less than two seconds, in which case two seconds is used as the interval. Each slide chosen is the one which has its PTS closest to (but not less than) the next frame interval 52. The last I frame in the video is generally selected as a slide, and an end-of-pass marker is associated with the last frame 54. Accordingly, a slide show representation of a 5 minute (300 second) video would include approximately 150 selected I frames.

10 Each selected I frame is marked with a second PTS 56 corresponding to its order and timing within the slide show. The frames are then stored in a temporary file according to their original order. The second PTS makes it possible to vary when and how long each frame is displayed during the slide show.

15 Once the frames comprising the slide show have been selected, the revised order of frames is stored in a temporary video file and indexed 58. As will be discussed in detail below, the slide show can then be viewed frame-by-frame 60, 62 by an operator using the video player component of the frame selector module 116. The video player utilizes standard MPEG video and audio decoders and has a  
20 rewind and replay function. The frame selector module 116 permits the operator to edit the slide show by adding or deleting frames 64 and 66, or by substituting individual frames 68 in place of ones picked randomly by the frame selector module 116. The frame selector 116 also allows the operator to add, delete or change slide show PTS values 70 in order to vary when and how long a slide is displayed. When  
25 the operator finishes editing, the frame selector begins 92 to write the actual RAV file which will be stored at the server site.

An RAV file header sequence is prepared containing information on the total number of video and audio frames in the video and the bit rate the download order  
30 was prepared for. The header sequence is encoded at the front end of the RAV file 94. The information blocks representing the I frames chosen in the first pass and the corresponding LBR audio (the entire LBR audio data stream) are written into the  
/ 6

front end of the RAV file 112 immediately following the header sequence 96. The file is written such that a portion of the LBR audio data precedes the initial corresponding I frame data. The remaining audio data is arranged in temporal order with the remaining I frame data, however, the file is written such that an audio  
5 frame is always downloaded sometime prior to its corresponding video frame. In this manner, LBR audio data is always available to be played when the corresponding slides are displayed. This addresses the experience that short gaps in audio playback are more easily discernable, and more distracting, than short gaps in the visual slideshow presentation.

10

On the second pass, the frame selector 116 selects video frames which, when played back with the video frames and audio data from the first pass, produce a low frame rate video ( $1/4$  to  $1/2$  the original frame rate) with LBR audio. This video plays back with good to very good motion. However, unlike the first pass, the  
15 second pass need not be downloaded in real time.

The frames on the second pass 74 are chosen in one of two ways, depending on the makeup of the MPEG file. If the total number of I frames in the file is more than 25% of all frames 76, then approximately every fourth frame is chosen 78  
20 (unless that frame was already selected during the slide show pass). If the fourth frame is not an I frame, then the next valid frame is chosen instead. If the number of I frames is less than 25% of the total number of frames, then the second pass consists of all the remaining I frames plus all P frames 80. This results in a video which displays at approximately  $1/2$  the original frame rate. The actual frame rate  
25 ultimately achieved depends on the combination of frames used to make the original video but can be from 5 frames per second (fps) to 15 fps. The quantity of data selected for the second pass is typically more than is able to be downloaded in real time over a 28.8 kilobaud modem connection.

30 The information blocks representing the video frame data chosen in the second pass are written into the RAV file immediately following the slide show video frame data and LBR audio 98.

A third pass 86 includes all remaining video frames which have not been selected in either of the two preceding passes. The information blocks representing the video frame data chosen in the third pass are written into the RAV file immediately following the video frame data chosen in the second pass 100. Like the  
5 second pass, the third pass comprises a quantity of data which is typically more than is able to be downloaded in real time over a 28.8 kilobaud modem connection.

In a fourth pass, the information blocks representing the MPEG audio data stream are written into the end of the RAV file, followed by the end of sequence  
10 block 102 and 104. Like the second and third passes, the fourth pass comprises a quantity of data which is typically more than is able to be downloaded in real time over a 28.8 kilobaud modem connection.

As discussed, the second, third, and fourth passes may have more data than  
15 can be downloaded in real time. Accordingly, the transfer can take place in the background without user intervention. For example, if a user is using the invention in the context of browsing the World Wide Web, a certain Web page might contain a video clip. The user, by actuating a software control, can choose to receive the video clip, which is then displayed as a slide show in a portion of the Web page. If  
20 the user decides to download subsequent passes, the user can continue to browse other Web pages as the download continues. When the download pass is complete, the user is alerted and given the option to return to the Web page containing the video to view the downloaded file.

By way of example, the previously described RAV file is arranged to  
25 download over a 28.8 kilobaud channel in the following order: slide show frames and low bit rate audio in the first pass, video frames for building a low frame rate video in the second pass, the remaining frames (frames for building the original MPEG video) in the third pass, and the high quality MPEG audio in the fourth pass.  
30 Given this arrangement, the slide show data and low bit rate audio would be downloaded or passed down first, so a slide show could be displayed during the download process. During the second and subsequent passes, the slide show and

low bit rate audio, or a higher quality presentation if more data is available, is shown during the beginning of the download. After the playback is finished, the download is able to proceed in the background until the pass is completed.

5           If the user terminal is connected to the network by a faster connection, more bandwidth is available to transmit more video data. In this case, the content provider might elect to arrange the RAV file so that video frames necessary to display a low-frame-rate video could be downloaded or passed down first, at the same time as the low bit rate audio data. In this way, a low-frame-rate video with  
10 LBR audio, instead of a slide show, can be displayed during the initial download process.

          In the latter case, video frames which would normally be selected in a first and second pass would be selected in a first pass for incorporation into the front end  
15 of the RAV file. The RAV file components would then be arranged to download in the following preferred order: video frames for building the low frame rate video, the remaining video frames, and the MPEG audio. The LBR audio is preferably downloaded simultaneously with the low frame rate video; alternatively, it can be downloaded before or after any of the RAV file components.

20

          It is also possible to assemble RAV files comprising a variety of different download arrangements including arrangements where audio data is downloaded last, or not at all, in which case a slide show or video could be displayed without audio. In this case, more frame data can be transmitted during the download  
25 process. These embodiments are also within the scope of the invention.

          The audio and video information blocks of the RAV file 112 would be prearranged in the necessary download order for a given baud rate and stored in that order at the content provider's server sites as a data structure encoded on a  
30 computer-readable medium.

#### Receive Sequencing Interface

When a client requests a video, a URL (Uniform Resource Locator) describing the name and address of the file to be downloaded is transmitted to the server 126 storing the RAV file 112. The server 126 uses the URL address to locate the RAV file 112. The server then forwards a URL to the Receive  
5 Sequencing Interface (RSI) 72 requesting authorization to begin transmitting the RAV file 112.

With reference to Fig. 3, the RSI 72 comprises a URL processor 130, a block transfer interface 128, frame builder 116, an index file generator 134, a frame  
10 sequence table 140, an audio and video playlist 136, 138 and an MPEG video decoder/player 144. The components of the RSI 72 cooperate to receive and process video data at the user terminal so it can be displayed.

Upon notification of the URL processor 130, the RSI 72 establishes a  
15 TCP/IP connection to the server via the block transfer interface 128 which starts a flow of block data from the server 126 to the RSI 72. The frame builder 116 stores the blocks of data in the order received so that the RAV file 112 is reassembled. At the same time, the index file generator begins to construct the audio and video playlists 136 and 138 and the frame sequence table 140.

20

The frame sequence table 140 is constructed from information extracted from the header of each video information block. The frame sequence table 140 has an entry for each block of video frame data. The layout of each entry is shown in Table C. The information in the frame sequence table 140 is used by the system  
25 sequencer 142 of the player 144 to locate video information blocks in the RAV file 112.

TABLE C

Field	Description
5 Item Type	This describe the kind of item that is in the Video File It could be a Header, A Group of Frames, Start of Frame, or End of Frame, etc.
Frame Type	If it is a frame, this will describe the type I, P, or B
Decode Order	This is the sequence number of the frame in the file. Frames are presented out of order because some frames need certain past and future frames for decoding
Display Order	This is the sequential Frame Number that would be shown to the user
PTS	Presentation Time Stamp for this item, if it is a displayable frame
10 Slide TS (Second PTS)	Presentation Time Stamp for this slide's appearance if it is a displayable frame in reordered slide show
Local File Location	The location of the Item in the RAV file, in bytes
File Location	Location in the actual video file
Slide Translation	How we go from one frame to another
15 Size	Size of the item

Frame Sequence Table

20 The video and audio playlists 138 and 136 are computed from information extracted from the RAV file header sequence. Each playlist consists of a plurality of entries, and each entry stores data for an individual video or audio frame. Each playlist is created with enough entries to accept information for every frame in the data stream. The information stored in an entry in the playlist is shown in Table D.

25

TABLE D

Field	Description
Index	A positive number is an Index into the Frame Sequence Table N negative number is a command to the system sequencer: -1 Skip this frame -2 Sleep and retry frame
30 Replay Number	Every time the user replays the video, more and more frames are available to view, the replay number tells you what replay sequence this slide is part of

Video Playlist



Field	Description
Index	An Index into the RAV File in bytes

5

### Audio Playlist

The video playlist 138 tells the system sequencer 142 in what order the video frames are to be decoded in a given cycle. The video playlist also contains pointers into the frame sequence table 140 for each frame entry.

10

### Frame Builder Module

Fig. 6 shows the operation of the frame builder 116 and index file generator 134. Initially, the video playlist 138 is created by the file generator 134 with a -2 in each index entry. As the first video frame to be played is downloaded 168, the frame sequence table 140 is updated 170, 172 with the location of that video frame block in the RAV file 112, and the negative number in the playlist 138 index entry corresponding to that video frame is updated with a positive number 172 pointing into the frame sequence table 140.

20

When the next video frame is downloaded 176, the frame sequence table 140 is updated 178 and the negative number in the video playlist 138 index entry corresponding to that frame is updated with a positive number 180 into the frame sequence table 140. The negative two (-2) in each entry between the entries containing the positive numbers is then changed to negative one (-1) 182, and the video data block is saved to the RAV file by the frame builder 184. As each video frame is downloaded, the process is repeated 186 until a positive number is entered in the video playlist 138 for every frame in the slide show and all of the intervening entries have been changed from -2 to -1.

30

The audio playlist 136 contains pointers into the RAV file 112. There are two audio playlists, the first list will be a pointer into the LBR audio. The second list will be a pointer into the MPEG audio. Audio frames are stored by the frame builder into the RAV file in the same order they are received 166. As each LBR

audio block is received 162, the file generator registers the audio blocks file location in its corresponding entry in the LBR audio playlist 220.

Once the slide show frame data and LBR audio data have been downloaded,  
5 the video frame data selected in the second pass is downloaded 188. As each video frame is received, the frame sequence table 140 is updated 190, a positive number is entered 192 in the corresponding entry on the video playlist 138 and the slide show PTS is deselected for every preceding video frame. The video data is then saved to the RAV file 194 by the frame builder. This process is repeated for the video frame  
10 data which was selected on the third pass 196. After the third pass data is downloaded, the frame sequence table 140 would contain a complete record of all video frame data and the video playlist 138 would have a positive number in every entry. The order of frame data in the video playlist 138 reflects the same order in which data is presented by an MPEG encoder to an MPEG decoder. The  
15 presentation order is different than the display order.

Once the second and third pass video data is downloaded, the MPEG audio data is downloaded 198, timing information is extracted 200, each audio frame is registered 202 in an entry in the MPEG audio playlist 136, and the MPEG audio  
20 frame data is stored 202 in the RAV file.

### Video Player Module

With reference to Fig. 3, the video player module 144 is shown. The player  
25 operates as a standard MPEG decoder/player, as shown in Fig. 2, except the standard MPEG system decoder is replaced with a system sequencer 142. The system sequencer 142 is responsible for synchronizing and directing the playback of the audio/video streams and is invoked as soon as the frame builder module 116 begins to receive the RAV file 112 from the server 126.

30

The system sequencer 142 determines the next frame to decode by looking at the video and audio playlists 138, 136 and the current status of the video and audio  
23

output buffers 146, 148 of the player 144. When the system sequencer 142 reads through the video playlist 138, it will retrieve the corresponding video frame block for each positive entry it comes to and forward the blocks from the RAV file 112 to the video decoder 154 for decompression. However, if the system sequencer 142  
5 sees that the video output buffer 146 is full or the audio output buffer 148 is near empty, the system sequencer 142 will look to the audio playlist 136 to determine the next audio frame to decode and retrieve this audio block from the RAV file 112 for decoding.

10       The video player module 144 decompresses audio and video frame data in the order presented by the system sequencer 142. Once decompressed, the video frames are stored in the buffers and displayed in the order and for the length of time referenced by the slide show PTS.

15       If the user chooses to replay the downloaded data after the end of the slide show, the system sequencer 142 will read through the video playlist 138 again decoding the corresponding video blocks for each positive number it comes to. Since more video frames will have been downloaded, more video frames will be available for decompression and the resulting video image will be enhanced. If all  
20 of the video frames in the second pass have been downloaded, the system sequencer 142 will be able to direct the playback of the low frame rate video with sound.

In one embodiment, the system sequencer 142 is disabled from selecting video frames from the second or third pass for decoding until the last frame in that  
25 pass has been downloaded and the system sequencer 142 has read an end of pass marker. In that case, a display on the user terminal screen indicates when a given pass is downloaded and the user can elect to replay the slide show or wait until the download is complete.

30

Depending on the arrangement of video data in the RAV file 112 and the amount of data downloaded, the following non-limiting playback configurations are  
24

possible: slide shows, with or without LBR audio, where the video frame display rate is from 1 frame every 10 seconds up to 4 frames per second; a standard video at a preselected frame rate from as low as 4 fps to 24 fps, with or without LBR audio; an MPEG video (if all I, B, and P frames are downloaded) with or without LBR  
5 audio; and, if the full RAV file 112 has been downloaded, an MPEG video with MPEG audio can be assembled and played back. Standard videos with frame playback rates slower than 7.5 fps are within the scope of the invention but are not desirable due to the poor image quality.

## 10 EXAMPLE TWO

In this alternative embodiment, the RAV file 112 is created from an MPEG video as described in Example 1 and stored at a server site. However, the RAV file 112 does not have to be downloaded in its prearranged order. Instead, the server  
15 site is equipped with a frame sequencing interface (FSI) which can rearrange, in real time, the download order of the RAV file 112.

With reference to Fig. 7, a video distribution system is partitioned into a content management system 118 which comprises the transcoder 120 and frame  
20 selector 116 programs, an FSI 204 which is located on the video pump 126 (the principal storage unit for the RAV files and index files), a title manager 206 for processing video requests from the user terminal, and a client 208 which comprises the RSI 72 programming for receiving and displaying the RAV file 112 and the player 144.

25

The video distribution system operates as follows. The user registers for the video service via client/title manager interaction. This process compiles user hardware and software configuration, preferences, and password data.

30 The user, in interaction with the title manager 206, will select a video either from a video guide provided by the title manager or from a Web site. The title manager then selects a URL specifying the address of the video at an appropriate

video pump, and transmits it to the client 208. The client then requests this video by transmitting the URL to the video pump 204.

The FSI and video pump system 204 respond by providing this video to the client 208 in a format and frame rate selected by the client, or one which matches the hardware configuration (e.g. modem speed) of the particular user. If the modem speed will not support the download of a video, the user will receive a slide show with real-time LBR audio. As the amount of local video data increases during the downloading process as described above in connection with Example 1, low frame rate videos can be displayed with progressively enhanced quality. Upon completion of the download, the user will be able to view the full frame rate MPEG audio/video presentation.

With reference to Fig. 7, the MPEG video files are converted to RAV files 112 within the content management system 118 where the transcoder 120 and frame selector 116 programs reside. The transcoder 120 and frame selector 116 may perform the same function in the same way as described in Example 1. Thus, video data is selected in four passes and an RAV file 112 is created in which the video data is stored in the following order: slide show frames and LBR audio, low frame rate video frames, remaining video frame data, and MPEG audio.

At the same time the RAV file 112 is being assembled, a primary index file 122 (Table E) is created (see Fig. 5, step 114) which contains a record of the download order of information blocks in the RAV file 112 and information for locating each block in the RAV file or the original MPEG file. The primary index file is stored with the RAV file 112 at the server site 126.

**TABLE E**

Field	Description
Block download order	Block location in RAV File in bytes

The frame selection process performed by the frame selector module 116 in

Example 1 is repeated so that download sequences for different baud rates can be calculated. For instance, if the user has an ISDN connection, it may be possible to download sufficient data to play a low frame rate video with LBR audio during the download process instead of a slide show. In that case, the frame selector module  
5 116 would make a first pass and select all the frames necessary to make a low frame rate video (all the frames that were previously chosen in the first and second pass). The remaining video frame data would be selected on a second pass and the MPEG audio would be selected on a third and final pass as described in Example 1.

10 After the operator has finished making the final slide selection, the frame selector 116, instead of writing a new RAV file, creates a secondary index file 122 (Table E) which records the new download order and information about where the blocks are located in the original RAV file 112.

15 The secondary index files 122 are stored with the RAV file 112 and primary index file at the server site 126. A number of secondary indices would be prepared for a variety of different download arrangements and each index would contain pointers into the same RAV file 112. Thus only one large AV file need be stored, along with a number of small index files 122.

20 When a user attempts to download a RAV file using a low bit rate connection (e.g. a 28.8 kilobaud modem), the RAV file 112 created by the content manager 118 will be downloaded directly by the video pump 204. When a higher bit rate connection (e.g. ISDN) is used, the FSI 204 will resequence the RAV file  
25 112, according to the information in the primary and secondary index files, so that the most appropriate sequence is used. All of the frame selection and ordering calculations would have been made, in advance, in connection with the content manager 118, and stored in an appropriate secondary index file as discussed above.

30 With reference to Fig. 7, the content manager 118 is responsible for transferring the RAV file 112 and index files to the FSI/video pump storage unit 204, and upgrading the database of the title manager 206 to include the new video  
27

clip title. In a preferred embodiment, the title manager 206 and FSI/video pumps would be located at the head end in an Internet or intranet service provider facility or on an Internet backbone.

5           With reference to Fig. 7, the FSI video pump 204 comprises a transfer monitor 124, storage for the RAV and index files 112 and 122, and a block transfer system 210. When the server 126 receives a URL from the client for a particular video clip, the server creates a TCP/IP socket connection to the client's RSI 72. The URL contains both address information into the RAV file 112 and client  
10 information. The server 126 starts the transfer monitor 124 by passing the name of the file to be transferred and the connection speed of the user to the transfer monitor.

          The transfer monitor 124 searches the index files 122 for the secondary  
15 index that contains the download sequence for the given connection speed. The transfer monitor 124 then uses the index 122 to locate the appropriate information blocks in the RAV file 112, so they can be downloaded according to the download sequence recorded in that index.

20           In one embodiment, the FSI can respond to a user request for a particular RAV file format. For example, a user may elect to preview a slide show first, even though the connection speed may accommodate the download and real-time display of a low-frame-rate video. In this case, the transfer monitor would accept the request and search the index files for an index which contains a record of a  
25 download sequence which is front loaded with slide show video frames, such as the RAV file 112 described in example 1.

          The transfer monitor 124 then uses the secondary index to locate the appropriate information blocks in the RAV file 112, so they can be downloaded  
30 according to the download sequence recorded in that index 122.

          The output of the transfer monitor 124 comprises a series of information  
28

blocks from the RAV file 112, transmitted according to the download sequence recorded in the appropriate secondary index file 122. The referenced blocks are then forwarded via the video pump TCP/IP block transfer system 210 to the client's RSI 72.

5

Once the FSI/video pump 204 has delivered the data blocks to the RSI 72, the data is processed by the RSI 72 and video player as discussed in Example 1 and as shown in Figures 5 and 6.

10

While certain exemplary structures and operations have been described, the invention is not so limited, and its scope is to be determined according to the claims set forth below.



**WE CLAIM:**

1. A method for encoding and decoding an audio/video data file, comprising the steps of:
  - obtaining a digitized audio/video data file comprising a video data stream representing a sequence of video frames having an original order and a first timing mechanism for indicating the display order of the video frames;
  - decoding the audio/video data file into at least its component video data stream;
  - reordering the video frames in the video data stream;
  - assembling a reconfigured audio/video file comprising the reordered video frames; and
  - displaying the reordered video frames according to a second timing mechanism.
2. The method of claim 1, further comprising the step of rearranging the video frames according to the original order so the original audio/video data file can be displayed.
3. The method of claim 2, further comprising the steps of:
  - decoding the audio/video data file into its component audio data stream; and
  - compressing the audio data stream to produce a low bit rate audio data stream.
4. The method of claim 3, wherein the assembling step further comprises the step of incorporating the low bit rate audio stream into the reconfigured audio/video file.
5. The method of claim 4, wherein the assembling step further comprises the step of associating the second timing mechanism with the low bit rate audio data stream to synchronize the low bit rate audio data stream to the reordered video frames.

6. The method of claim 1, wherein the reordering step further comprises the steps of:
  - selecting video frames that can be used to assemble a slide show having a first pre-selected display rate;
  - associating the video frames in the slide show with the second timing mechanism to indicate the display order and timing of display of each slide show frame; and
  - selecting video frames which can be combined with the video frames selected in previous passes to assemble a video file having a second, higher, display rate.
7. The method of claim 6, further comprising the step of selecting all remaining video frames not previously selected.
8. The method of claim 7, further comprising the step of storing the video frames in a reconfigured audio/video file.
9. The method of claim 1, further comprising the steps of:
  - receiving and processing a request for the reconfigured audio/video file from a user terminal; and
  - downloading the reconfigured audio/video file from a storage unit to the user terminal in the selectively reordered download sequence.
10. The method of claim 9, wherein the download is accomplished in at least two passes.
11. The method of claim 10, wherein the download is accomplished in four passes.
12. The method of claim 11, further comprising the step of arranging the video frames for downloading in the following order: slide show frames, subsequent pass video frames, final pass video frames.

13. The method of claim 11, further comprising the step of arranging the reconfigured audio/video file for downloading in the following order: slide show frames, subsequent pass video frames, final pass video frames, original audio data stream.
14. The method of claim 10, further comprising the step of arranging the video frames for downloading in the following order: low frame rate video frames, subsequent pass video frames.
15. The method of claim 1, wherein the assembling step further comprises the step of creating an index file representative of an order for the reconfigured audio/video data file.
16. The method of claim 1, wherein preparing the reordering step comprises the steps of:
- using a frame selector module in at least two successive passes to select and identify appropriate video frames and their download order; and
  - creating and updating an index file with a timing and a sequence for each frame.
17. A system for encoding and decoding an original audio/video data file having an original order, comprising:
- a transcoder module for decoding the original audio/video data file into a video data stream;
  - a frame selector module for specifying a reordered sequence for the video data stream;
  - a storage unit comprising a computer readable medium, for storing a reconfigured audio/video file representative of the original audio/video data file in the reordered sequence; and
  - a user terminal through which a user may display the reconfigured audio/video file.

18. The system of claim 17, further comprising:  
a frame sequencing interface for retrieving and sending the reconfigured audio/video file from the storage unit to the user terminal; and  
a receive sequencing interface for downloading and processing the reconfigured audio/video file for display.
19. The system of claim 18, wherein the transcoder module is capable of decoding an original audio data stream from the original audio/video data file and compressing the original audio data stream into a low bit rate audio data stream.
20. The system of claim 19, wherein the transcoder module creates a plurality of data files representative of the video frames of the original audio/video data file, the original audio data stream, and the low bit rate audio data stream.
21. The system of claim 17, wherein the reconfigured audio/video file comprises an original audio data stream, a low bit rate audio data stream, a video data stream, and a first timing mechanism associated with each stream indicating a display order for the audio stream data and the video stream data.
22. The system of claim 21, wherein the video data stream of the reconfigured audio/video file comprises data representing a sequence of video frames.
23. The system of claim 18, wherein the video data stream of the reconfigured audio/video file is ordered so that video frames comprising a slide show are downloaded first.
24. The system of claim 23, wherein the reconfigured audio/video file is ordered so that the low bit rate audio stream is interleaved with video frames comprising the slide show, such that the slide show with low bit rate audio can be displayed while the data is being downloaded.
25. The system of claim 24, wherein the reconfigured audio/video file is

structured such that additional video frame data is downloaded after the data comprising the slide show and low bit rate audio.

26. The system of claim 25, wherein the reconfigured audio/video file is rearranged into its original order.

27. The system of claim 26, wherein the original audio/video data file has a video frame rate between approximately four frames per second and approximately thirty frames per second.

28. The system of claim 19, wherein the original audio data stream is downloaded after all video data stream information.

29. The system of claim 17, wherein the reconfigured audio/video file is arranged for playback according to its original order.

30. The system of claim 23, wherein each video frame in the slide show is associated with a second timing mechanism to indicate a display order and time.

31. The system of claim 30, wherein the slide show has a display rate between approximately one frame every ten seconds and approximately four frames per second.

32. The system of claim 31, wherein the display rate is approximately one frame every two seconds.

33. The system of claim 31, wherein the display rate is variable.

34. The system of claim 24, wherein the low bit rate audio stream has a transmission bandwidth between approximately twelve and approximately fourteen kilobits per second.

35. The system of claim 18, wherein the video data file is reordered so that video frames comprising a low frame rate video are downloaded first, so that the low frame rate video can be displayed while the data is being downloaded.

36. The system of claim 35, wherein the video data file is reordered so that the low bit rate audio stream is passed down at the same time as the video frames comprising the low frame rate video, such that a low frame rate video with low bit rate audio can be displayed while the data is being downloaded.

37. The system of claim 18, wherein the frame selector module prepares an index file representative of an order for the reconfigured audio/video file.

38. The system of claim 37, wherein the frame sequencing interface utilizes the index file to create a reconfigured audio/video file for downloading.

39. The system of claim 38, wherein the frame sequencing interface selects and downloads video data in an appropriate order selected from one of the following orders:

- a) slide show frames, low frame rate video frames, faster frame rate video frames;
- b) low frame rate video frames, faster frame rate video frames;
- c) slide show frames with low bit rate audio, low frame rate video frames, faster frame rate video frames, original audio;
- d), low frame rate video frames with low bit rate audio, faster frame rate video frames, original audio;

40. The system of claim 39, wherein the appropriate order is selected according to a download speed.

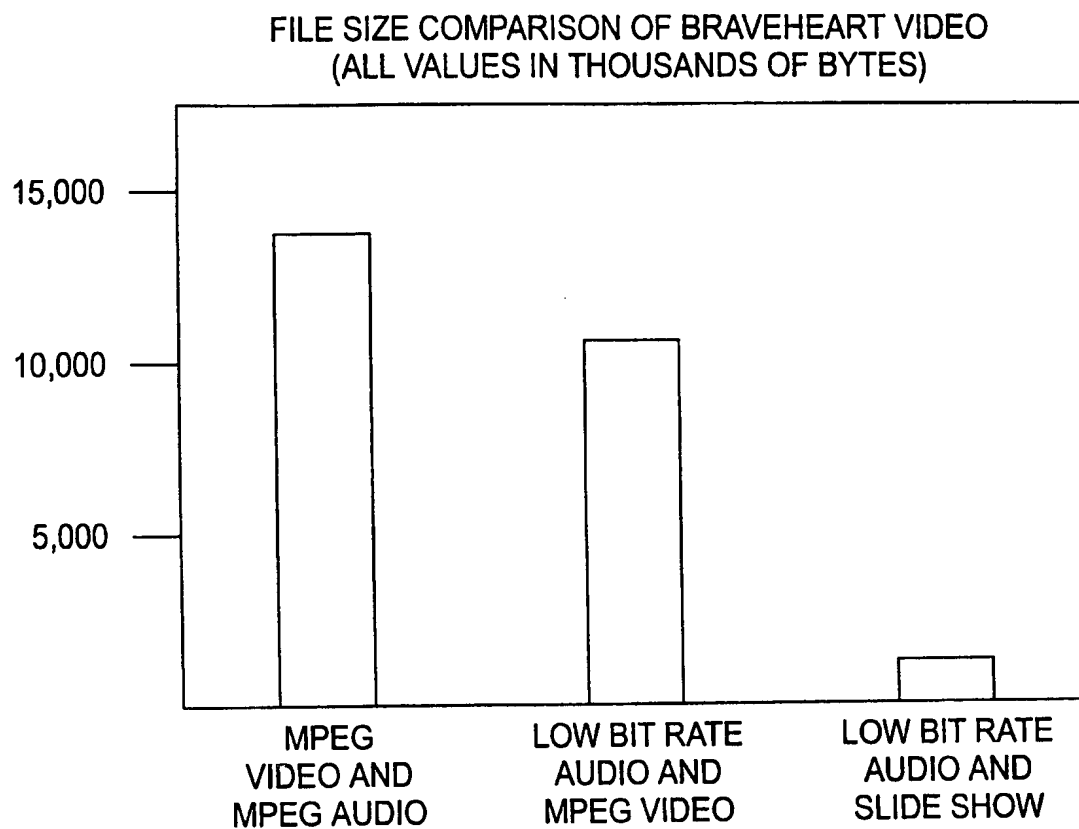


FIG. 1

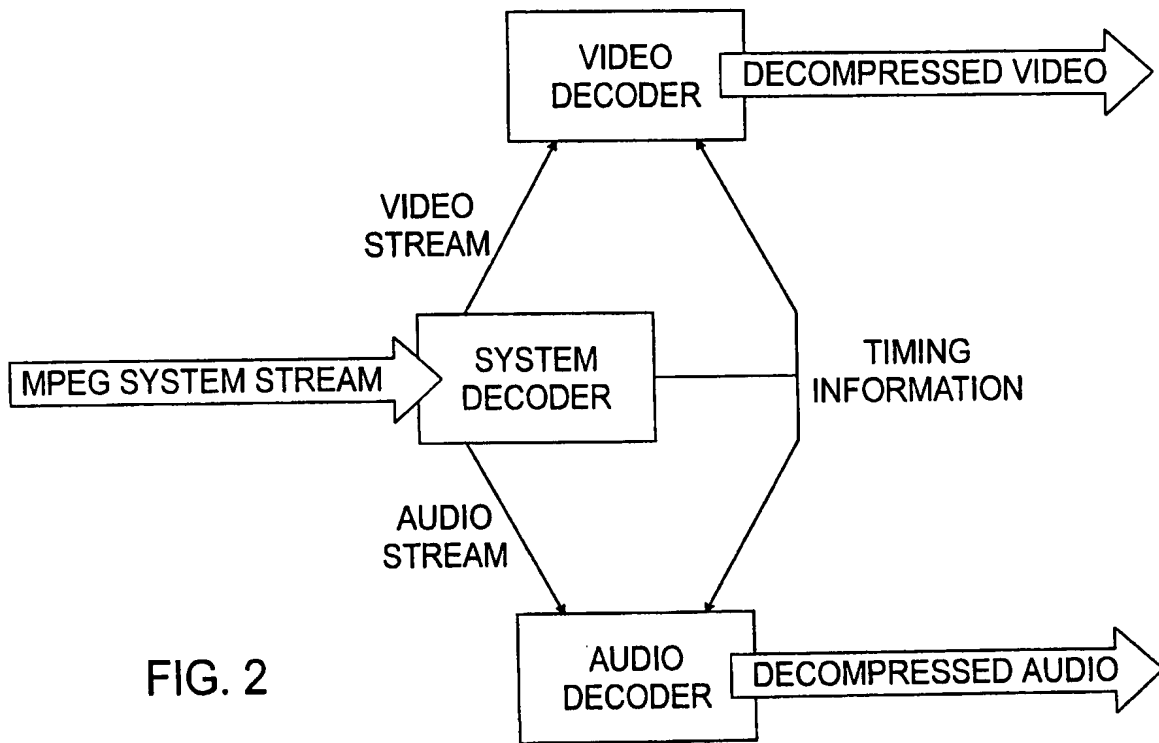


FIG. 2



FIG. 3

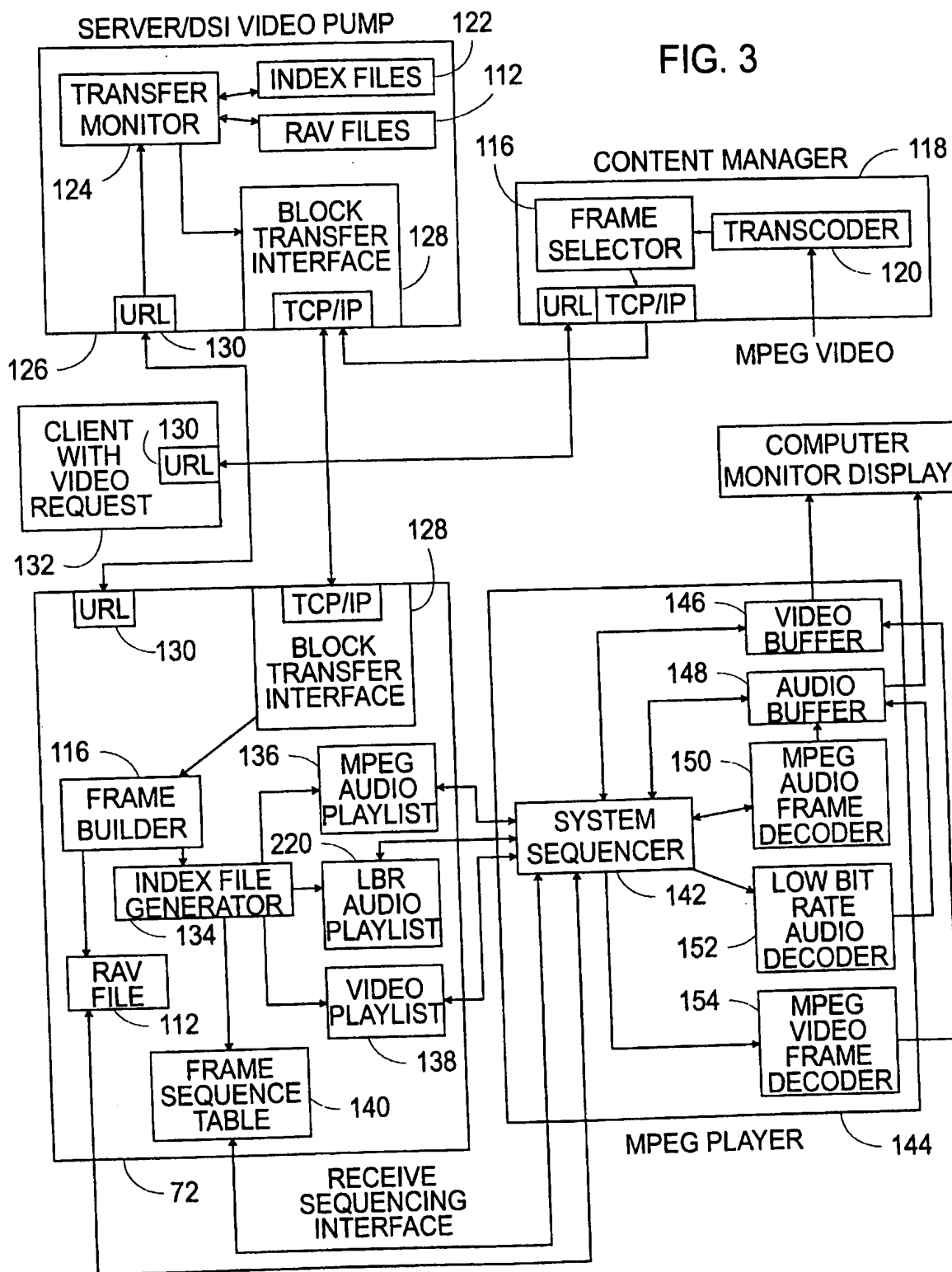


FIG. 4

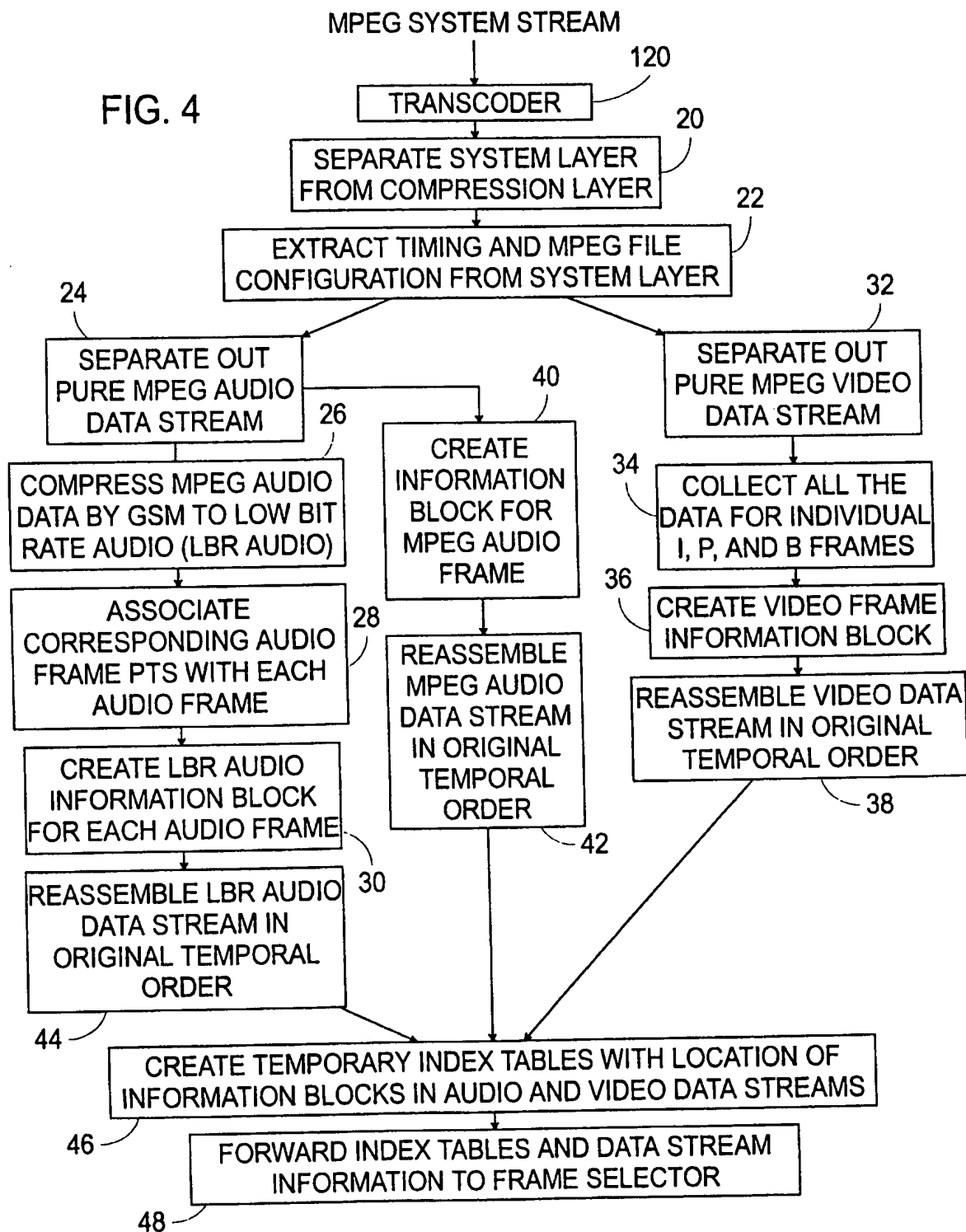
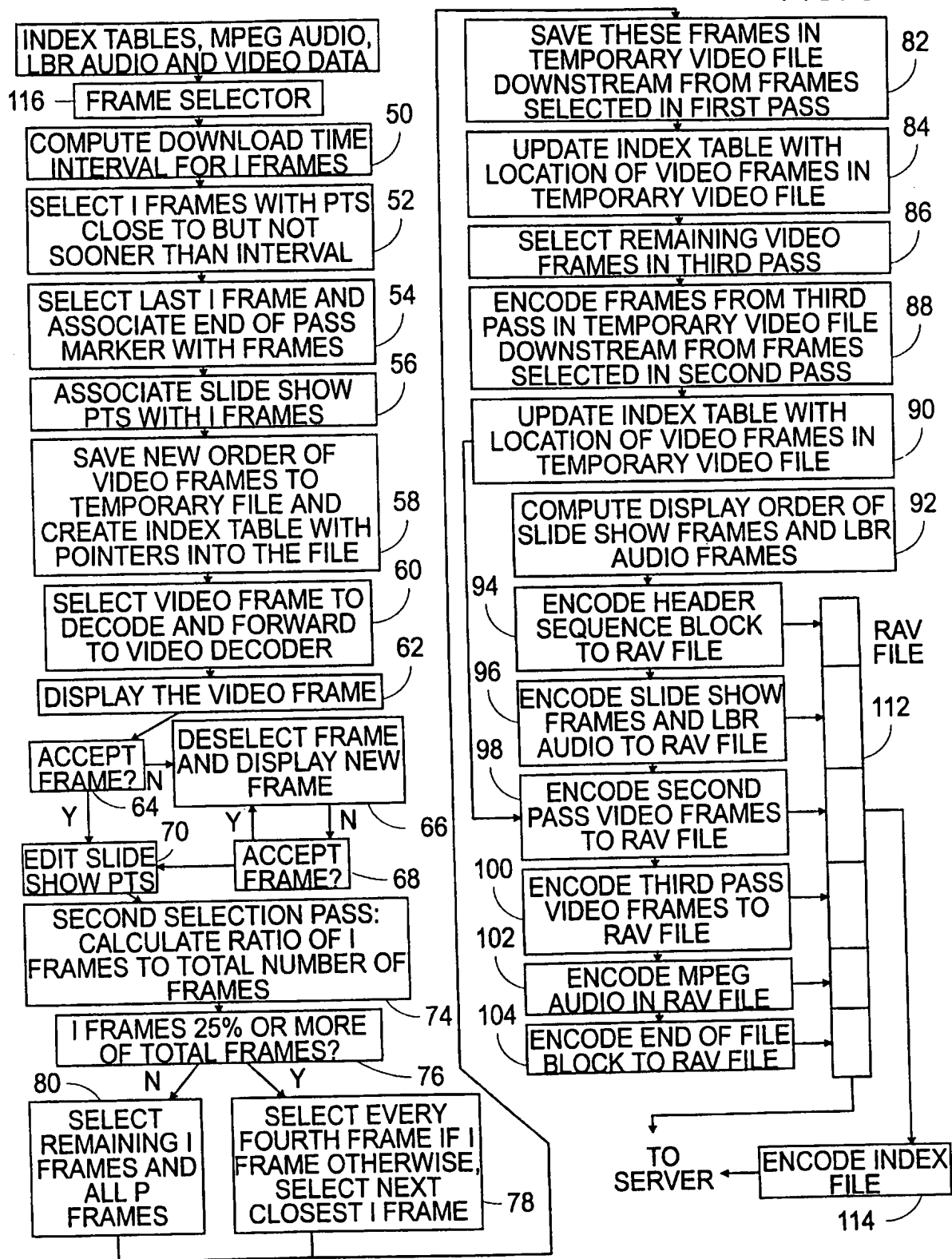


FIG. 5



6/7

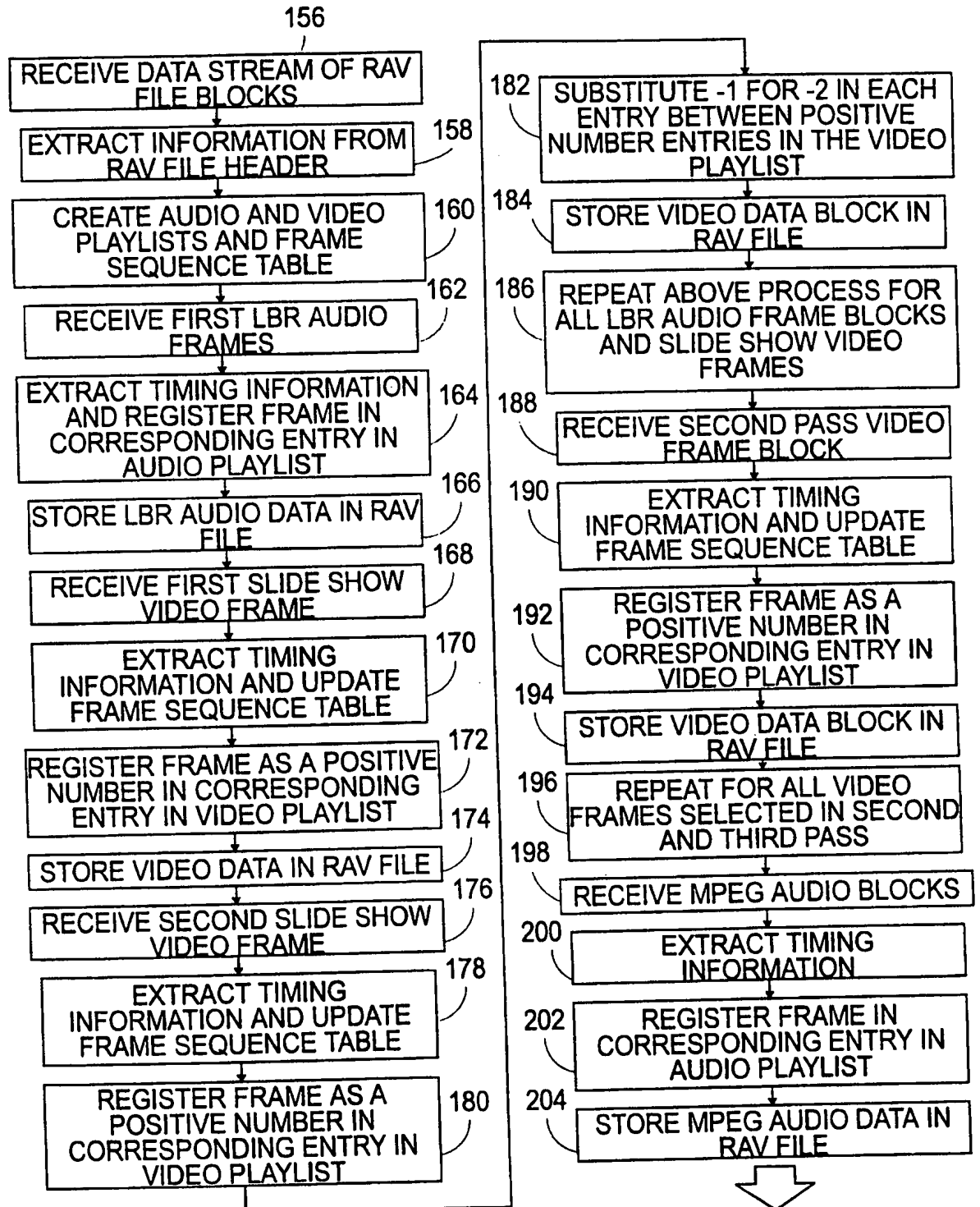


FIG. 6

RAV FILE, PLAYLISTS, AND FRAME SEQUENCE TABLE ARE QUERIED BY SYSTEM SEQUENCER

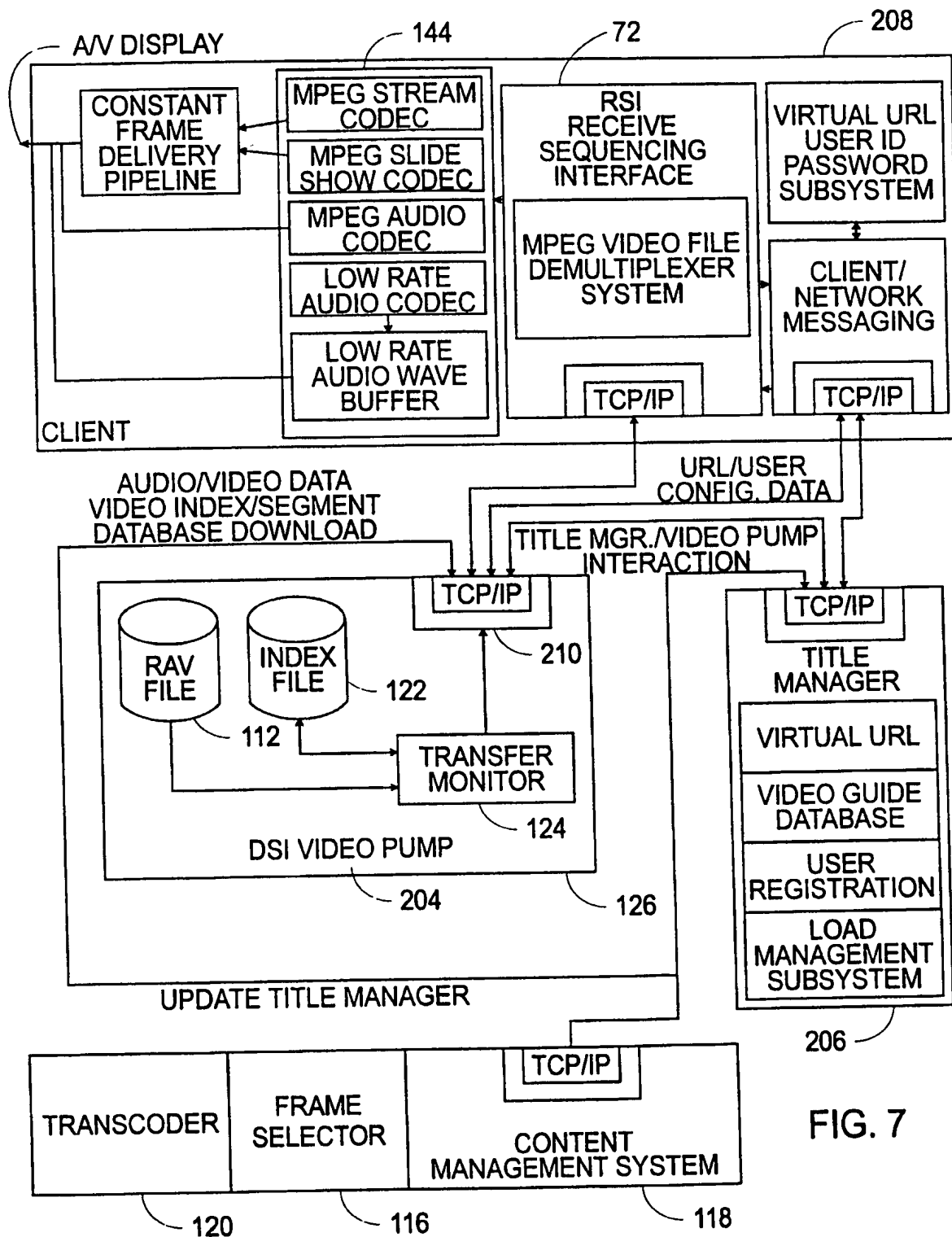


FIG. 7

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 98/03904

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 6 H04N7/26

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 737 930 A (SUN MICROSYSTEMS INC) 16 October 1996 see column 1, line 41 - column 2, line 40 see column 9, line 1 - column 12, line 31; figures 4-8 ---	1-40
X	EP 0 739 140 A (SUN MICROSYSTEMS INC) 23 October 1996 see page 2, line 11 - line 28; figure 1 see page 3, line 3 - line 14 see page 4, line 20 - line 23 see page 5, line 52 - line 56 see page 6, line 34 - line 47 see page 8, line 6 - line 54 see page 9, line 11 - line 17 ---	1-40
X	EP 0 702 493 A (IBM) 20 March 1996 see the whole document --- -/--	1,2,9

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance  
"E" earlier document but published on or after the international filing date  
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)  
"O" document referring to an oral disclosure, use, exhibition or other means  
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone  
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.  
"&" document member of the same patent family

Date of the actual completion of the international search

16 June 1998

Date of mailing of the international search report

24/06/1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Giannotti, P

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 98/03904

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 96 41285 A (INTERVU INC) 19 December 1996 see abstract see page 27, line 5 - line 13; figure 1 see page 49, line 10 - page 50, line 21 ---	1-40
A	EP 0 661 881 A (SHARP KK ;NIPPON TELEGRAPH & TELEPHONE (JP)) 5 July 1995 see abstract ---	1-40
A	EP 0 687 112 A (MATSUSHITA ELECTRIC IND CO LTD) 13 December 1995 see abstract ---	1,17
A	EP 0 726 677 A (NIPPON ELECTRIC CO) 14 August 1996 see abstract -----	1,17

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 98/03904

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
EP 0737930	A	16-10-1996	CA 2173698 A		13-10-1996
			JP 9167124 A		24-06-1997
EP 0739140	A	23-10-1996	US 5621660 A		15-04-1997
			JP 9163362 A		20-06-1997
EP 0702493	A	20-03-1996	JP 8098166 A		12-04-1996
			US 5742347 A		21-04-1998
WO 9641285	A	19-12-1996	AU 6113996 A		30-12-1996
			EP 0834143 A		08-04-1998
EP 0661881	A	05-07-1995	JP 7184176 A		21-07-1995
			JP 7184189 A		21-07-1995
			JP 7184177 A		21-07-1995
			JP 7184178 A		21-07-1995
			US 5594736 A		14-01-1997
EP 0687112	A	13-12-1995	JP 7336681 A		22-12-1995
			JP 8084339 A		26-03-1996
EP 0726677	A	14-08-1996	JP 8223556 A		30-08-1996
			AU 4337596 A		15-08-1996
			CA 2168531 A		10-08-1996